# 2010 CENSUS BUREAU PRESS BRIEFING
# NOVEMBER 1, 2010

**STAN ROLARK:** Good afternoon. My name is Stan Rolark. I'm Chief of the Census Bureau's Public Information Office. I want to thank everyone and welcome everyone for being here today, both those in the auditorium as well as those on the phone, and those of you who are looking at us online. We're very happy to have you here.

Today, Dr. Robert Groves, the Census Bureau's Director, will give an update of our 2010 census operation. And he'll also talk briefly about our 2020 census planning. For those of you who want additional information, you can go to our website, www.census.gov. You'll see an icon in the upper left-hand corner. If you click on that icon, you can get information about this press conference.

As always, after Dr. Groves speaks, we'll open it up to questions and answers. We'll alternate as we've done in the past between those in the auditorium here and those on the phone. I will ask that you keep it to one question. And certainly, if you have another question, you can get back into the loop. For those of you who are on the phone, but sure to cue up the operator. Let her know that you have a question and that you'd like to be put into the lineup to ask your question. So, without further adieu, I'd like to bring in Dr. Groves. Dr. Groves?

**DR. ROBERT GROVES:** Thank you, Stan. It's great to see you here, thanks for coming. Today, I'm going to give you an update on the Census Bureau's progress in completing the 23rd census of population that we all participated in a few months ago. I can say in that regard that we are well on our way to deliver the results back to the society that produced them in the beginning, and we're looking forward to that.

I'm going to give you some updates on evaluations of census quality, as much as we have them now. We're just gathering these proxy indicators now. Say a bit about the data processing steps that we are going through right now, which eventually produce the

statistics. And then as Stan mentioned, I want to say a few things, believe it or not, about the 2020 census and plans that we are considering for that.

So let me jump in. As I've said before, one of the key evaluative vehicles for any census is sometimes called the post enumeration survey. It's a very large sample survey that's done after the census data collection is complete. And we go out to a sample of households and we do an in depth interview that determines whether we correctly captured that household in the census. Through that process, we identify misses, people that were missed in the census. We identify erroneous enumerations from time to time, people that shouldn't have been counted that were. And we estimate those successfully counted.

We have just completed all of the interviewing for this decade's post enumeration survey which we call the Census Coverage Measurement Operation. In a nutshell, things went well. Let me give you some statistics on that. We had a big interviewing force, as we did in the census itself. We had re-interviews of their work to check whether they were following training guidelines. 99.7 percent of those interviewers passed that re-interviewing check. Another way of saying that, we only had 18 interviewers that failed that check. And as we did in the non-response follow-up operation, their work was completely redone when we discovered that. So we had a great staff, they did their work quite well and we have some other production data off of that.

In the 2000 census at this point, finishing that post enumeration survey, about 5.3 percent of the cases produced information on the population that came from a neighbor, or a building manager. It wasn't a self report, it was what we call a proxy report. This time, we did better. That same number is 3.8 percent, a little lower, and we're proud of that. On the other hand, and a negative signal, is in the year 2000 about 0.14 percent of the cases when we finished all of our efforts we still didn't have a population count on, it was a non-interview case. This time, that 0.14 has risen to 1.54 percent. So that's something on the other side of things.

We've also been able to match all of the cases in this big sample survey to our big census master address file. So an immediately interesting question is does it match up well? Does it look like we have a good census master address file? And there, the news is all good, no matter how you look at it. This year, we're estimating at this point that about 96.5 percent of the addresses match up to the master address file that we used to mail out all the cases. Last time, in 2000, that 96.5 percent number was 91.4. Similarly, 96 percent of the cases we judge were correctly enumerated. Based on that match, compared to about 89.9 percent in 2000. We also have fewer duplicates in this check.

So what does this mean? The big picture here is all of these are comments, initial indicators, about whether our address list was good, was strong. Did we miss addresses? And it looks like we're in better shape than we were in 2000 on that. Let me move to something else. As you may recall, when I came in, I thought it prudent to introduce into this census a test of internet technology. We have done that. We labeled this census quality survey. What we want to do with this test is to see whether people answered questions on the internet in a different way than they do on a paper questionnaire. This is really critical, looking forward to 2020 on how we should use the internet, what sort of population groups are best suited to it.

We've completed that survey. A sample of households around the country will compare their answers that they sent in from the census to what they got on the survey. We'll have a report on this early next summer, I think. We have a few other evaluations, believe it or not, that are still going on, but these are minor things.

Let me turn to big operational issues. We had 494 local census offices. We're closing those down in a very careful manner. We've closed more than 59 percent of them at this point. As of this morning, that's 293. We think we'll close all of those by November 12. This is not just kind of locking the doors and walking away. We have computer networks in these offices. We have a team that goes in and completely sanitizes the computer, the desktops, the Xerox machines. We want to make sure every trace of confidential information is wiped off these machines before they're moved out of there.

We've also rented furniture. Vendors come to pick that up. All of the material that is owned by us is returned to our central processing unit in Jeffersonville, Indiana, and then we inspect it and it will eventually be disposed of through normal GSA procedures. I want to remind you that we also have 12 big regional census offices. Those are going to stay open a little longer to deal with personnel issues that linger after you hire a million people; and our 12 permanent regional offices will remain.

What we're doing in Suitland right now is a very different task than we've been doing over the last couple years. We don't have a big physical presence in the nation, we have people who are scrutinizing gigantic data files and making sure that we understand the responses that people intended to give us and that we correctly capture them. And I want to give you some statistics on that. We process the paper forms that we receive either through a mail back or through an interviewer administration with optical scanning machines. These are really cool machines. The technology for optical character recognition has become more and more sophisticated. We enjoy the benefits of that sophistication.

We have captured 3.3 billion fields of information, 3.3 billion fields. These are the little checkboxes that you did on your own form. And as you might imagine, a lot of things happen when you look at those 3.3 billion. Sometimes, people don't check nicely through the box. Their X or their checkmark extends down to an adjacent box. What's been great about the developments in optical character recognition is that we're able to read those more and more successfully.

So for those kind of checkbox fields, we failed to read about .1 percent of those in 2000. This time, we failed to read .03 percent. Now, what does that mean? When the machine senses that it's not confident that it understands what people checked, we move that to a human keyers. One of our staff members examines the digital image of the checkbox and then make a decision. That's a lot of money we spend on humans. If we can reduce that, it's good, and we've done it. You also may remember in some fields that the respondent

actually wrote in an answer. This was true on race and ethnicity, especially. These machines are reading those characters and the rate of successful reads has gone up, I mean do the opposite. In the year 2000, we failed to read about 16 percent of those written-in answers. This time, we're failing to read about 12 percent. It's an improvement on that side.

So these sound like small differences. But when you're processing hundreds of millions of forms, we now can say that we've saved 27 percent of the staff costs because we didn't have to move those machine reads into a review step, and that's about $36 million. So we've saved $36 million on processing the 2010 census forms, mainly because of these advances in technology.

In your press kit, you have a two-page little document that is labeled “Advances due to the capture of data in the margins of questionnaires.” And for the first time-- this is really a geeky kind of thing but I've got to tell you about it-- for the first time, these machines are sensing writing outside of the boxes, in the margins. And we're trying to use that information wherever we can, as much as we can. So you're looking for something that looks like this. It has some blue boxes, two blue boxes on the front page. So one of them shows you an example of something that happens more often than one might expect. A household with a little baby writes in to answer the question for age the number four. But outside the box, they write in months. They clearly want to tell us that this is a four month old baby and we can capture that with this new technology to make sure we don’t interpret this as a four-year old baby, that it’s really a four-month-old baby. We're proud of these advances and we think this is going to improve the quality of the data.

You have another handout that gives you a description of one other big thing we're doing. It has a bunch of rectangles with numbers in it. So you might want to look at that. Let me tell you how to read this stuff. Imagine that each of those rows of data described a different household. And so, you go way up to the top on the left, there is a household that has an I.D. code of 12694. And all of those numbers in the unedited data file are

essentially answers to questions that the household gave us. And there are particular things that I want you to know that we're attending to.

First, notice that about six down, there appears to be a duplicate I.D. Sometimes, people fill out the mail questionnaire, mail it back, but they mail it back too late for us to register it and we send an interviewer out, and we completely redo that data record. We have duplicates now. That's an example of a duplicate. So one of the things we're doing right now with these big data files is resolving duplicates. In the year 2000, maybe about 10 percent of the file had duplicate I.D.s.

Then notice that there's some big black squares on the data file, the unedited data file. Those are missing data. Sometimes, people don't answer a question that they should have answered. We have to deal with those. And then if you look every once in a while, there's a light gray number. And that's there to distinguish or to convey a written-in answer and we have to code those. We're doing all that work right now. We are resolving duplicates, we are dealing with missing data, we are coding data and that involves a whole lot of people and a lot of subject matter expertise to make sure we do it correctly.

The other thing we're doing, and we will release this later this month, is resolving all the geographical changes. Those who use census data for redistricting need to know the boundaries of blocks and tracks, block groups, in an unambiguous way so that when they form congressional district boundaries, they have clear sub units with clear geographical boundaries. We are in the middle of doing that. We'll release these state by state on a flow basis. November 23rd is a target date. And then we are stripping off from the files that I just mentioned all personally identifiable information. So we are going to break the lock between names and addresses and the data themselves in producing these files to make sure we keep the pledge of confidentiality that we've given everyone.

We're on track, there's a real tight schedule, but we are on track. These things are going well. We are on track to release the state protein totals by December 31st, which we're

mandated to do. And shortly thereafter, we will begin state by state to release the files, the data files that are used for redistricting, giving data down to much smaller areas. So, so far, so good on that.

I want to stop, this is all kind of retrospective talks; now let's look forward. December has three big releases. They're going to hit boom, boom, boom. The first release will be a release of what we call demographic analysis population estimates. These will be national totals broken down by age, gender, and black/non black status. These are not 2010 census data. These are not 2010 census data. Let me say it one more time. These are not 2010 census date. These do attempt to estimate the population of 2010, but it uses a very different technique. It uses birth registration data, death registration data, estimates of immigration and emigration, in and out migration. We've used this, we used this throughout the decade. We update these constantly. Every developed country in the world has such a way of estimating national totals.

In the past, indeed, demographic analysis has been used to evaluate how good the census is. We are less and less comfortable doing that. In January of 2010, we brought together a great group of demographers around the country to tackle the problem of how we estimate components of immigration and migration. It became clear that there was no consensus on how to do this well. There are rival estimates. Each of the rival estimates has a rationale that is plausible. None of the rationales beat out others. And so, for the very first time we're doing two things. We are releasing the demographic analysis early, earlier than the census. This is something we can do because of the American Community Survey. And we're not going to give one national total for each age, gender, black/non black group, we're going to give multiple ones to reflect our uncertainty about the actual totals. This is the only honest thing we believe we can do. It will indeed be a little more complicated for people who want the certain answer of what is the number, but it will be more honest.

Next in December, we will release the five year period estimates for the American Community Survey that reflect the years 2005 through 2009. These will be very rich data

describing characteristics related to social issues, education, marital status, and so on; economic issues, income, employment, occupation, and housing and demographic characteristics. They'll be released for the first time for areas with less than 20,000 people. 670,000 geographical areas will get their own estimates. These are not 2010 census data. These are from a sample survey called the American Community Survey. This is a wonderful thing that we are giving back to the country to help guide local decision making for businesses and local governments. And they will describe a period of time that has passed. And so, there's another handout that you have that looks sort of like a PowerPoint. And I want to give you a sense so that you can convey it to your readers and listeners about what these five-year period estimates might look like.

On the first page, in red, is portrayed in a dashed red line, is portrayed-- it looks sort of like this. Hopefully you have it in your press kit. The red line attempts to portray what the five year period estimate might be for a phenomenon that is described by the blue line. So, if you could imagine, these are all made up data, really just to convey a point. In this first one, we're estimating the percent 25 years old and over with a bachelors degree. And notice that in 2005, we pretend we knew this at 20 percent, and it goes up slightly at 21 percent. In fact, it’s going up slightly each year over the five year period. A five year period estimate for that kind of phenomenon averages out those five years in the following way. You see the red versus the blue.

The important point to note for all of us who absorb these is it won’t look often-- the data we release there won’t often look like what you see outside of your window. The world has changed since this period 2005 through 2009. And we're giving the country back some notion of an average over the five year period. If you turn the page and look at another example, a volatile kind of phenomenon, home ownership rates, and these are all made up data again, something that's going up and down, you again see how the red line, the ACS number, is going to be kind of an average of those five numbers.

And finally, if you have something that isn't changing at all, the very last one, percent in management, professional related occupations, the five year will be identical to the

unchanging single five year estimates. So again, when we release this, this will be a ton of data. It'll be done on very small areas. There are wonderful stories to be written about our area, our city, but they need to be written in a way that it's clear the reader is learning about the city as it was 2005 through 2009. It's not the 2010 phenomenon, but something else.

And then finally at the end of December, finally, finally, we will release the state population totals for the country. And since we are charged with the burden, the honor of the arithmetic for calculating the reapportionment of the House, we will also release at that same moment the number of House members by state.

So now let me turn way forward and say a few things about the 2020 census. Believe it or not, in the midst of all this, it is our obligation to look forward. And we are doing this with our senior team at Census and we have articulated a set of principles that will guide 2020 planning, or should guide it, in our belief. And the very first is that we are attempting to design a 2020 census that costs less per housing unit than the 2010 census while maintaining the quality of the results. We are serious about costs at the Census Bureau, reducing them, because we believe that if you examine the increase of costs of censuses over our lifetimes, we are on an unsustainable trend. We can't continue the trend that we are now on.

I remind us all that at least in my lifetime, this country went through a fundamental shift in census taking. It was between sort of the 1950s and 1970s, really 1960 to '70 where instead of having enumerators go out to every house in the country, we mailed. That was a cost reducing change, a fundamental change. We are comfortable with that change now. Many of us think that's what a census is. It is time, we believe, to sit down and rethink the fundamentals of a decennial census and we're doing that. And we're doing that actually for the same reason it was done those 40 or 50 years ago, depending on how you count. And that is we have reached the end of cost sustainable designs in how we're doing it.

If you look at the cost components of a decennial census, it's really easy to find a big component to attack. That component is the salary cost and infrastructure, supervisory support, for the large numbers of people who knock on doors. If we want to save money in the next census, we have to reduce the number of human resources that are needed to do it. It's that simple. If we want to do that, we have to somehow improve self reporting, or find ways of getting data, suitable census data, on people without knocking on their door.

So we have committed to what could be called a multiple mode census. The current census is a multiple mode census. It uses a mail questionnaire and a follow-up with a face to face interview. We believe other modes have to be there. We can commit to an internet option in 2020. We will commit to other modes as they make sense. One can imagine a variety of electronic communications that will be available in 2020. We can imagine the use of the phone as other countries have done.

There are also principles of the planning we've enunciated. One, we have committed to a faster cycling of ideas, smaller tests, a lot of small tests versus a small number of very large, expensive tests. Two, we're committing to using our existing infrastructure. We're going to try in all ways not to invent one time infrastructure to do the testing, but to use existing infrastructure. Three, a great example of that is the American Community Survey. So, we want to use the American Community Survey as NASA used the shuttle for the last many years, adding experiments onto that vehicle to learn things, to learn new things, to try out new ideas in a pretty rapid way.

Fourth, we are attempting to avoid a software development philosophy that says, "You will develop a complicated system of software that must work perfectly the first time you use it in production, and then you throw it away." This has been the software philosophy of the decennial census for many, many years. Instead, again, we have an opportunity to use the American Community Survey as a production environment to do initial tests of 2020 software, and then ramp it up when we need it.

And then fifth, this is somewhat redundant, but we won't examine ways to improve the quality of the census without simultaneously looking at their cost impacts on the final census. I need to note, there's a special case in options that we're looking at. One option that's being used increasingly around the world to reduce the burden of the census is to use records that people have already filled out reporting the same information to some other government agency. I have a lot of emails in my box that say, "Why are you asking us questions that we've already answered already to another government agency?" We want to study whether this is acceptable both technically and legally and socially in this country. And so in the years ahead, we'll be looking at various administrator record systems collected by the federal government to see if we could reduce the burden of data collection by leveraging that information.

So those are my remarks. Next month, we're not going to have a press briefing. But as I just said, we're going to give you a lot of stuff. We'll send out the demographic analysis results at a meeting that's open to you. We will release the five year American Community Survey 2005 through 2009. That'll be open to you. And then finally, the big event at the end of the year where we release the state totals. I hope you're there at that time. After the holidays, we begin the redistricting data, rolling it out state by state in a way that serves states needs. But that's our December looking forward. So thank you, and I'm open to questions.

**STAN ROLARK:** Thank you, Dr. Groves. Now, before we open for questions, let me just remind you of just a couple of things. If you're on the phone and you have a question, remember to let the operator know so the operator can cue you in. So if you have a question when you call up, let the operator know that you want to ask a question.

Also, Dr. Groves had quite a few handouts. And for all of you who want to channel your inner geekdom, you can go to our website, www.census.gov and you'll see the icon in the upper left-hand corner so you can click on that and you can see those handouts. So again, we'll ask questions. We'll alternate from in the room and then online. So let's start with a question in the room. Carol?

**CAROL MORELLO:** Carol Morello with the *Washington Post*. Dr. Groves, you said, if I heard you correctly, there has been a tenfold increase in the percentage of people who you had no records at all on from .14 percent to about 1 ½ percent? What is that in raw numbers? And why is that happening, and what are the implications?

**DR. ROBERT GROVES:** Okay, let's make sure we're talking about the same thing. The numbers I quoted today pertain to the census coverage measurement program, the vehicle that we use to ask the question, "How good is the census?" So one of the results of that large sample survey is a larger, non-interview rate than we had in 2000. The numbers you cited were about right. It's a tenfold, that's one way to look at it. The other way to look at it is it's 1.5 percent versus .014 percent.

To the extent that that is an undesirable result, you're correct, the extent to which that harms this evaluative tool of the census really depends on how well we know the characteristics of those households and then can adjust for that non-observation. But the most important thing, I think, on this is to make sure that we all understand that's on this evaluative tool, not on the census itself.

**STAN ROLARK:** Thank you. And operator, we'd like to go to the phone lines now. We have a question, please?

**OPERATOR:** Yes, we have a question from Hope Yen at the Associated Press. Your line is open.

**HOPE YEN:** Hi there. I actually wanted to follow up on Carol's question a little bit. I was wondering if those numbers relate to a potential undercount that we've seen traditionally in previous censuses?

**DR. ROBERT GROVES:** They do. I hear you, Hope, asking the question, "Do these numbers relate to the traditional undercount?" They are relatable. How they relate is

really kind of complicated. If I ask that question, I can't do a direct comment on the proxy rate and the non-interview rate that we were talking about, how that pertains to the eventual undercount statistics from this. It is a weakening of that evaluative tool, I'd have to say that. But it doesn't mean the tool end up estimating bigger or smaller undercounts.

On the other hand, there's good news out of there, too, that I view as much more relevant to the eventual undercount statistics. And that is this match rate between the listings we did on the sample survey and the big master address file. And those match rates look better this time. Again, it's early and so on. But that's in the direction of saying, "Gee, we might have missed fewer houses than we did in 2000." That's good news, if that hangs in there.

**HOPE YEN:** Could you then, based on all the numbers, provide kind of a big picture assessment on how confident you feel the census will be able to improve the undercount compared to previous times?

**DR. ROBERT GROVES:** You know, I didn't hear all of the question, Hope. Could you do it one more time?

**HOPE YEN:** Yes. I was just wondering, based on the numbers you cited and just explained, what is your sense, your big picture sense, of how confident you are that the census will be able to improve on some of the undercounts we've seen in previous times?

**DR. ROBERT GROVES:** This I take, then, as really a question about the quality of the census itself; not the quality of the evaluative tool of the census, but the quality of the census itself. Well, I go back and say, "Gee. Look, we have initial indirect evidence on the whole household count that we may be in better shape." That doesn't say anything about missing people within households. But, so far, it looks good relative to 2000. It is way too early, however, to answer your question. I don't have a professional judgment at this point, and I don't have a level of confidence because we're way at the beginning of this process. And the results of the census coverage measurement program, this post-

enumeration survey, will not really be available until 2012. That's when I can answer that a little better.

**HOPE YEN:** Okay, thanks.

**STAN ROLARK:** Thank you. Let me remind everyone, too, when you have a question to please give your name and your media affiliation before you ask your question. So now we'll move to the room. Do we have a question in the room?

**AL MILLIKEN:** Al Milliken, AM Media. I have a friend who was working in the Commerce Department and recently got transferred to Suitland to the Census Bureau. I was wondering how many new full time census people have recently come on? And what will they be doing from what they had been doing, do you know?

**DR. ROBERT GROVES:** I don't know the answer to this. It's great that your friend has joined our little family. But this is not the time of growth of the U.S. Census Bureau. We are declining, if anything. We've used a large number of contract employees during this ramp up to the decennial and as their work is finished, they are leaving us. It's a bittersweet time at the Census Bureau. This is not a growth industry right now.

**STAN ROLARK:** Okay, let me remind everyone as well that you can always call the Public Information Office if you need more information. It's 301-763-3691. So operator, we're going to move to the phones. Do we have another question on the phone, please?

**OPERATOR:** Yes, we have a question from Richard Massey, Arkansas Democratic Gazette. Your line is open.

**RICHARD MASSEY:** Hey, Dr. Groves. Thanks for having this meeting today. Does the Census have an estimate if the State of Arkansas will gain any seats in the U.S. House of Representatives after 2010?

**DR. ROBERT GROVES:** You know, it's a good question and it's a question not about Arkansas, but name a state. It's a question that's often asked of me. And I got to tell you something that I don't know whether I should be embarrassed about this or not, but I don't follow this at all. I am the worst person to ask this question. I have no idea. I don't even read these newspaper articles. So, my job and the job of my colleagues, is to get the counts right. Then we do a little arithmetic to determine the numbers. But we're not forecasting that at all. I have no idea how that's going to turn out. But I do promise you we'll do our best to get an accurate count of Arkansas and do the arithmetic right to determine the counts. That's about all I can do.

**STAN ROLARK:** Okay, thank you for that question. We have one in the room now. Sabrina, you had your hand up?

**SABRINA TAVERNISE:** I'm Sabrina Tavernise from the *New York Times*. Just a follow-up, I'm sorry to be thick, but what precisely is the 1.5 percent versus 0.14 percent referring-- what is that measuring? And then my second question is if you have a sense of what the response was by undocumented immigrants? I don't know if that's how you kind of think about that?

**DR. ROBERT GROVES:** I apologize for being obtuse on this. I think we ought to go back a step and ask the question, "How could you ever answer the question of how good the census is?" A traditional tool that's used around the world is to do a completely independent data collection of the census. Draw a sample of areas, and in those areas use really extensive methods, the best we can think of. Go in and we talk to sample households, long questionnaire. The questionnaire's all about where you were on April 1 and did anybody move in or out? And then we do that completely independently of the census. In fact, the people that do it, believe it or not, for the last year, have been in locked rooms inside our office. They're told not to talk to anybody with regard to the census. This is like an independent evaluation.

Then we get those data from the sample survey and we match it up to the census and we say, “How many people have we found? Did we find more people in these houses than they reported on the census?” From that, we get the kind of estimates that we all really want to know right now, and that is what was the undercount? Is it differential across subgroups?

On that sample survey this decade, we have this non-interview rate that we've been talking about. And we have some proxy rates. The proxy rate looks better than it was in 2000. The non-interview rate looks less good, it’s worse than it was in 2000. So that we report to you because that's sort of an evaluation of the evaluation tool itself. We're way out there, right? That should not in itself reflect what the eventual differential undercount would be. It just means that our tool to estimate the differential undercount has these properties, some of which are better, some of which are worse. And we wanted you to know that. We’ll keep reporting this kind of evaluation of the evaluation as much as possible so everybody can make judgments.

When we eventually put out the differential undercounts, you should ask the question, “How good was that thing that produced the differential undercounts? I mean, why should we believe these differential undercounts,” right? And that's a good thing, and we’ll try to be as open as we can. Am I speaking to your question? This is complicated stuff. This is not unlike the method that's used in fisheries to answer the question, “How many fish are in this stream?” So what do you do in that? You grab a fish, you go in and you sample a fish and you put a little tag on his gill and you throw it back in. And you grab another fish, and you keep doing this repeatedly. Pretty soon, you start grabbing fish with things in their gills, right? And you say, “Gee, this is the same fish I sampled last time,” as a way to estimate the population. This, in essence, is sort of what we're doing on a much bigger scale.

**SABRINA TAVERNISE:** ... (inaudible)

**DR. ROBERT GROVES:** Everybody's in this. We won't be able to separate-- when all is said and done, since in this survey we don't ask citizenship and we don't ask documentation status, we won't be able to break the differential undercount by documentation status.

**STAN ROLARK:** Thank you for that question. So we'll next take another question in the room. Yes, sir?

**ROBERT THOMASON:** Thank you. Robert Thomason, Global Resources News. I just wanted to ask about plans for presenting the data to the public, if there are going to be any alterations, improvements, new challenges? Particularly with geo coding and mapping? I'm not talking about redistricting, but generalized?

**DR. ROBERT GROVES:** We are really pumped about some things you haven't seen yet. We have redone some software that you and your family and just plain folks can use to access census data. It's called the American Fact Finder. I don't know whether we have videos up on our website yet? In January, we'll have some videos, kind of tutorials on how to use it. It is greatly enhanced. We're hoping that it's user-friendly in a threshold change way. And we can't wait to get your feedback on whether it does the trick on geographical presentation. And also, ease of just putting queries in that the public is interested in asking. So we're working on that.

**STAN ROLARK:** Okay, thanks, Robert. Questions inside? Carol?

**CAROL MORELLO:** Dr. Groves, I know it's just a small grain of sand on a big beach. But I'm sort of curious about those 18 enumerators, who you said their work had to be completely redone. Can you tell us a little bit about some of the reasons why, how much it cost to redo the work? Were they all in one area? Are any of them being prosecuted? Can you give us more information?

**DR. ROBERT GROVES:** I can’t give a lot, unfortunately. But we're happy to share it with you if you want to drill down on this. With the number 18, there's no particular clustering, I think I can give you that, for each of the 18. The reasons that arise, that can arise, are really variable. But it comes down to a concern about whether all of the training guidelines, especially the crucial ones for quality, were found. It doesn't automatically mean that there was clear falsification of data. Sometimes, it was sloppiness to the extent that we couldn't accept their data and they were dismissed and we redid the work. But we could get you-- I don’t want to fake it. We’ll get you the information if we want. We have to protect the identity of these staffers, but we can give you the reasons, I think.

**STAN ROLARK:** So we have a caller on the phone. Operator, can we have the question, please?

**OPERATOR:** Yes. Hope Yen, the Associated Press. Your line is open.

**HOPE YEN:** Yes, hi it’s Hope Yen with the AP again. A question regarding the demographic analysis survey you're talking about. I just wanted to confirm that you said there will be multiple national estimates put out and that this will probably not be a good predictor of the 2010 census numbers. I guess I'm just wondering what we can draw from all the numbers if there ends up to be wide variation between the demographic analysis numbers and the 2010 count?

**DR. ROBERT GROVES:** I haven't seen the numbers, the numbers don’t exist, first of all. When we reveal them, we’ll all be seeing them about the same time. As a statistician, as an official statistical agency, we have to keep faith with the people that when we put out data that we can defend them. It was our professional judgment that putting out a single estimate for these demographic analysis statistics wasn't the best we could do. So when we revealed these, we’ll also describe the sources of the variability and why it’s plausible to believe one versus the other.

I must admit that means your stories are going to be a little more complicated, and how to write them-- we're happy to be completely transparent on this. But if we did the opposite, if we chose one of those numbers, we couldn't feel good about them professionally at this point. This country is observing phenomenon that make it very difficult based on these kinds of data sources to estimate the total population. The census is the official count, and we sought out direct measurements of all groups. But this relying on record systems means that you tend to miss people that fall out of record systems. And one of those has to do with the immigration statistics.

So, it's a horrible answer to your question, Hope, I know. But in our professional judgment, we couldn't justify a single estimate, that the first thing. And secondly, we believe this will be an honest presentation of alternatives. And talking about those is actually good for the country, I think, to talk about different estimates and why they exist. We'll learn something as a society about this, but it's a tough problem.

**HOPE YEN:** And so you're saying that you don't think these numbers will offer a good predictor of what the 2010 census count numbers will be?

**DR. ROBERT GROVES:** Well, they're going to offer multiple predictors. There are going to be multiple national counts and they're all plausible. Each of them has a defender. You will hear some of the defenders when we present this, and they'll give you the rationale why their number is the correct one. But you have multiple numbers out there. So, matching those up to the state totals and the national total a few weeks later is itself a rather complicated job. But that's the state of human knowledge of the population count these days.

**STAN ROLARK:** Thank you for that question. Do we have any questions in the room? Sir? Mic will be there in just a second.

**AL MILLIKEN:** Al Milliken again. Before the final count is certified, if any individual believes they were not counted, is there anything more they or anyone else can do at this point in the process to try to still be counted?

**DR. ROBERT GROVES:** Well, they should call a number. This is very, very, very late, right? We processed every form by September the 8th. But for the relief of someone who is worried about this, that number could be called. The chances at this point in time of including that return are miniscule. It is very late in the process.

**AL MILLIKEN:** The specific example that I know about-- actually, I handed this person a census form earlier, they didn't fill it out. They are working on a part time basis, they do not have a home, haven't for some time. When they can afford it, they stay in a hotel. But I think the night of April 1st, they were in an all-night diner and I don't think there are any efforts to check that specific location for an individual like him. And I don't think he particularly cares whether he was counted or not. But it is a concern to me.

**DR. ROBERT GROVES:** This is a complicated case, and we know and we must admit, my suspicion is, for evermore in this country, those folks who don't have a permanent residence, who weren't in the shelters that we had enumerated prior to the census and then visited, weren't served by the soup kitchens on the day that we were-- weren't living in an outdoor location that we had pre-designated working with local community groups, those people we missed. And I suspect we will always miss such people. And every democracy has that problem in doing a census. If people actively evade that measurement, we're not good at counting them. For those without permanent homes, it's a disproportionate challenge in this country, and every country I know.

**STAN ROLARK:** Okay, we have about two more minutes. This will be the final question. Are there any other questions in the room? We have Carol.

**CAROL MORELLO:** Dr. Groves, is there anything new to report, if there are any decisions been made, are you close to making any decisions on future counts of same sex

couples and that question that was rather controversial on race this year involving whether to continue to use the word Negro?

**DR. ROBERT GROVES:** We have, as you know, or hopefully you know, there will be a special tabulation released using the current measurement process in the 2010 sample for same sex couples where one of the members of the couple is person one, will have that relationship. That'll be done in May, I think. There is a group studying the measurement properties on this. We've reached out to communities that have given us great input on how to do this, some experimentation is going on. So, on the same sex measurement problem for sample surveys in the coming years, there may be some innovation that we'll put out for public comment.

On race and ethnicity, it is my judgment, and I've talked with the chief statistician at OMB, that this decade we need to do a serious review of that measurement. It is, as you know, under the authority of the Office of Management and Budget that sets the exact categories of race and ethnicity, and they will exercise that authority this decade as well. We want to work with them in doing the right kind of research and outreach to the diverse groups in the country to do that well.

When to do that in the decade is an interesting problem. If we perfectly solved race ethnicity measurement for 2011, that doesn't mean it will work well in 2020. And so getting the words right, getting the right groups measured specifically needs to be timed in a way so that it still works in 2020 and we have to worry about that. So we haven't launched these activities yet. We're talking about when to launch them and how to launch them at this point.

**STAN ROLARK:** And with that, I'd like to thank everyone for attending today. Thank you, Dr. Groves. Let me remind you as well, if you want an electronic copy of the press kit, go to www.census.gov. And if you have questions, follow-up questions, you can call the Public Information Office at 301-763-3691. Thank you.

END OF PRESS CONFERENCE